Higher-Order Factor Models

- Topics:
  - The Big Picture
  - Identification of higher-order models
  - Measurement models for method effects
  - Equivalent models

Sequence of Steps in CFA or IFA

1. Specify your **measurement model**(s)
   - How many factors/thetas, which items load on which factors, and whether your need any method factors or error covariances
   - For models with large numbers of items, you should start by modeling each factor in its own analysis to make sure *each* factor fits its items

2. Assess model fit, per factor, when possible (if 4+ indicators)
   - **Global model fit**: Does a one-factor model adequately fit each set of indicators thought to measure the same latent construct?
   - **Local model fit**: Are any of the covariance discrepancies problematic? Any items not loading well (or are too redundant) that you might drop?
   - **Reliability/Info**: Are your standardized loadings practically meaningful?

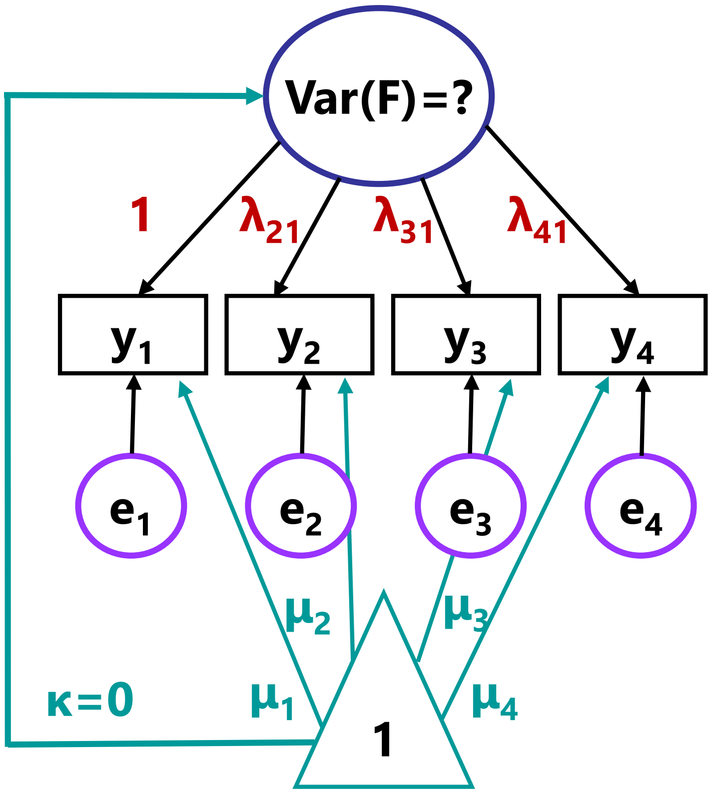
3. Once your single-factor measurement models are good, it's time to consider the (higher-order) structural model

Higher-Order Factor Models

- Purpose: What kind of higher-order factor structure best accounts for the **covariance among the measurement model *factors* (not items)**?
  - In other words, what should the **structural model among the factors** look like?
  - Best-fitting baseline for the structural model has all possible covariances among the lower-order measurement model factors → **saturated structural model**
  - Just as the purpose of the measurement model factors is to predict covariance among the items, the **purpose of the higher-order factors is to predict covariance among the measurement model factors themselves**
  - **A single higher-order factor** would be suggested by **similar magnitude of correlations** across the measurement model lower-order factors

- Note that distinctions between CFA, IFA, and other measurement models for different item types are no longer relevant for the higher-order model!
  - Factors and thetas are all **multivariate normal latent variables**, so a higher-order model is like a CFA regardless of the measurement model for the items
  - Latent variables don't have means apart from their items, so those are irrelevant

Necessary Measurement Model Scaling to fit Higher-Order Factors



"Marker Item" for factor loadings
→ Fix 1 item loading to 1
→ **Estimate** factor variance

Because it will become "factor variance leftover" = "disturbance", factor variance **can't be fixed** (it must be estimated)

"Z-Score" for item intercepts or thresholds
→ Fix factor mean to 0
→ Estimate all intercepts/thresholds

All the factor means will be 0 and you generally won't need to deal with them in the structural model anyway

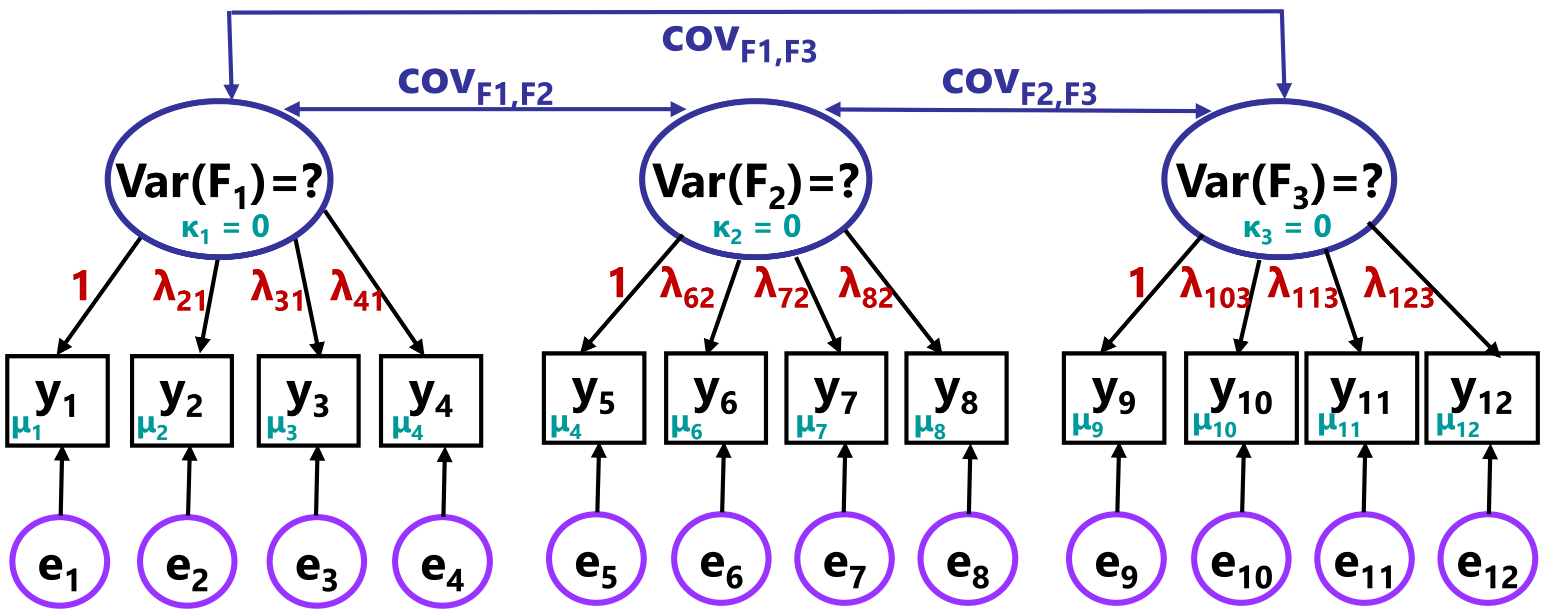
Identifying a 3-Factor Structural Model **Option 1**: 3 Correlated Factors

Measurement Model for <u>Items</u>:
item variances, covariances, and means

Possible DF = (12*13) / 2 + 12 = 90
Estimated DF = $9\lambda_i + 12\mu_i + 12\sigma^2_{e_i} = 33$
DF = 90 – 33 = 57 → over-identified

Structural Model for <u>Factors</u>:
factor variances and covariances, no means

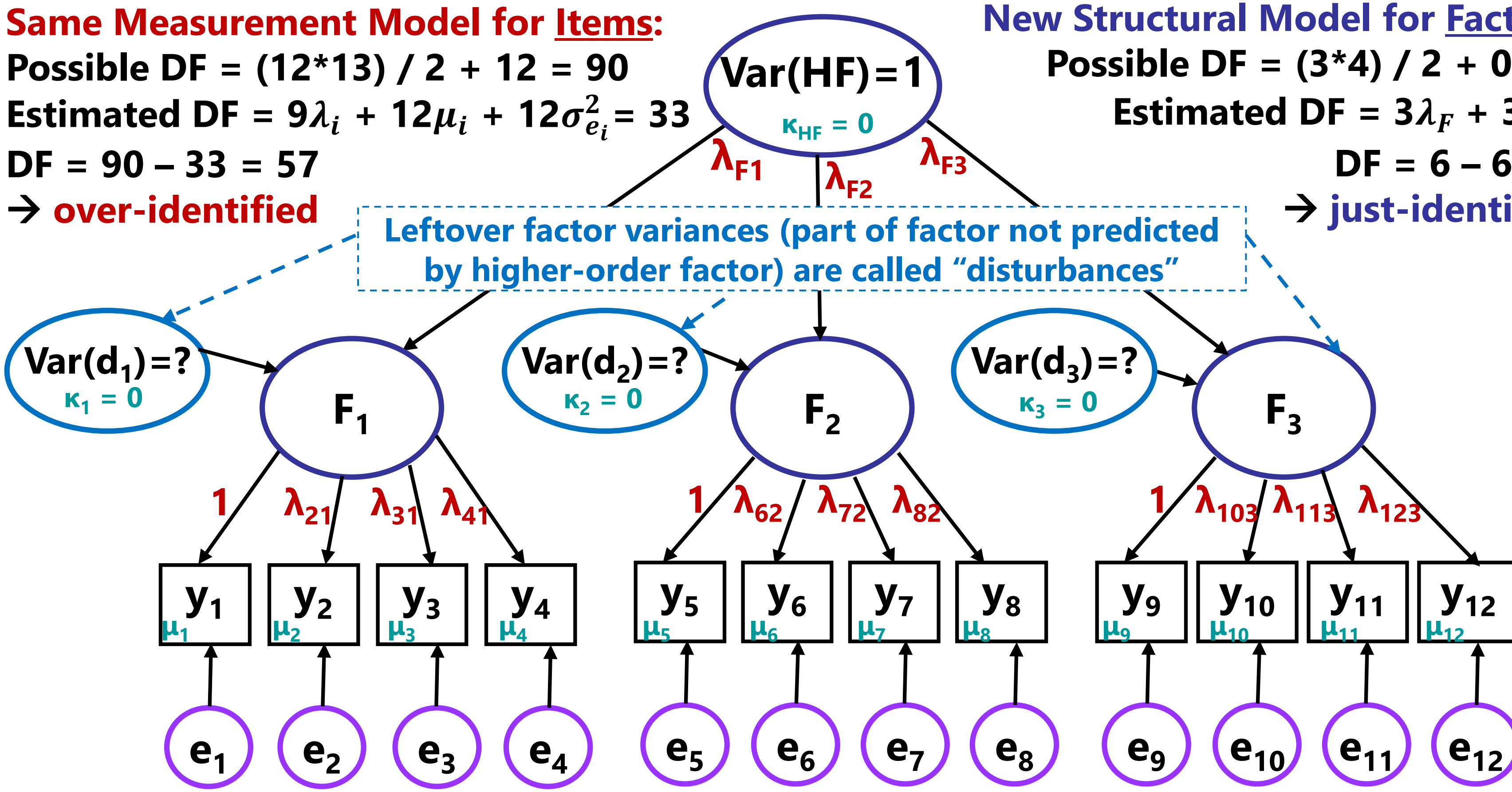
Possible DF = (3*4) / 2 + 0 = 6
Estimated DF = 3 variances + 3 covariances
DF = 6 – 6 = 0 → just-identified



Option 2a: 3 Factor "Indicators" (Higher-Order Factor Variance = 1)

Same Measurement Model for <u>Items</u>:
Possible DF = (12*13) / 2 + 12 = 90
Estimated DF = $9\lambda_i + 12\mu_i + 12\sigma^2_{e_i} = 33$
DF = 90 – 33 = 57
→ over-identified

New Structural Model for <u>Factors</u>:
Possible DF = (3*4) / 2 + 0 = 6
Estimated DF = $3\lambda_F + 3\sigma^2_{d_F}$
DF = 6 – 6 = 0
→ just-identified

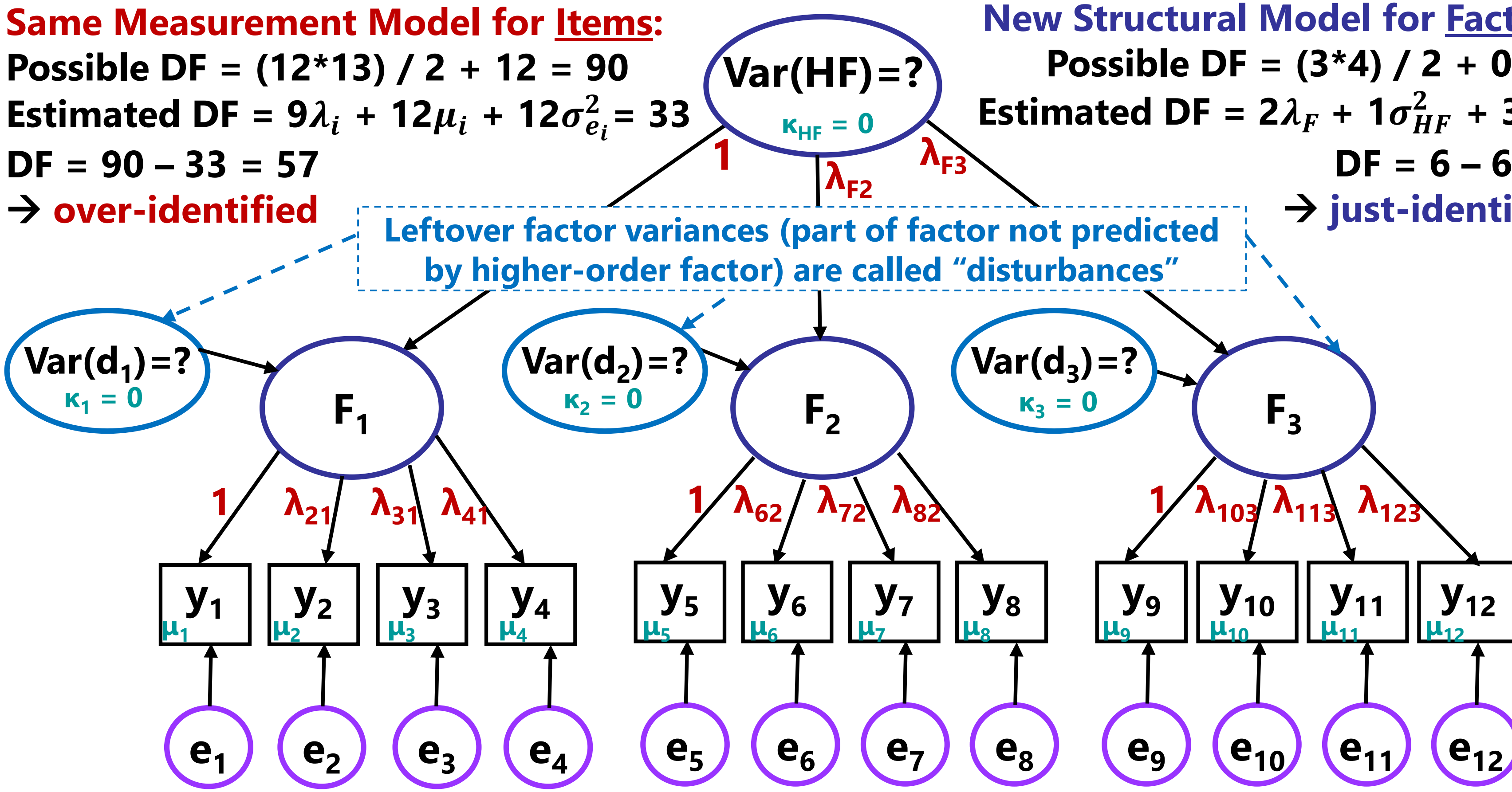


If you only have 3 factors, both models will fit the same—the structural model is just-identified, and thus <u>the fit of a higher-order factor CANNOT be tested</u>

Option 2b: 3 Factor "Indicators" (using Marker Lower-Order Factor)

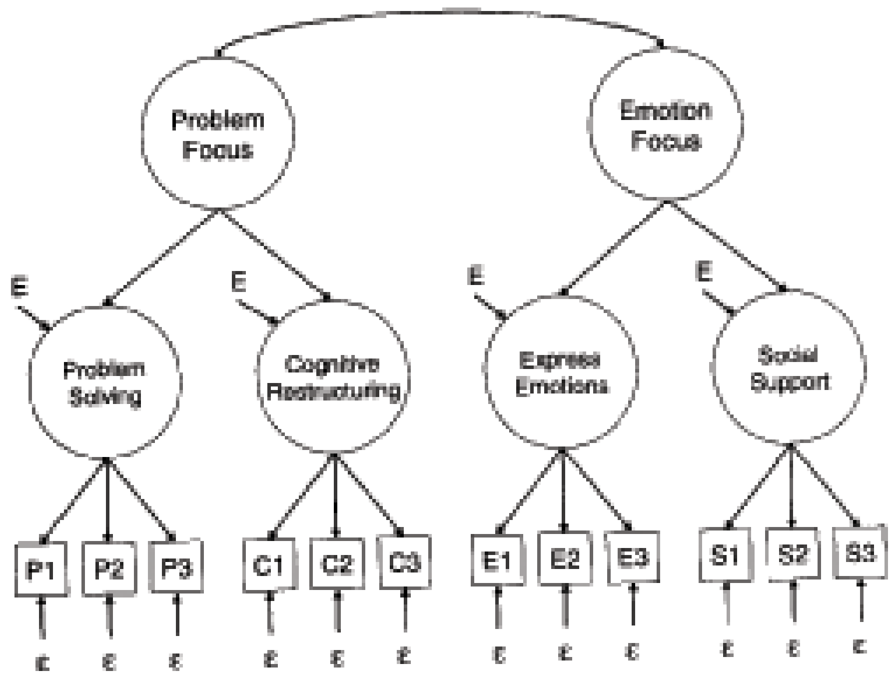
Same Measurement Model for Items:
Possible DF = (12*13) / 2 + 12 = 90
Estimated DF = $9\lambda_i + 12\mu_i + 12\sigma^2_{e_i} = 33$
DF = 90 – 33 = 57
→ over-identified

New Structural Model for Factors:
Possible DF = (3*4) / 2 + 0 = 6
Estimated DF = $2\lambda_F + 1\sigma^2_{HF} + 3\sigma^2_{d_F}$
DF = 6 – 6 = 0
→ just-identified



If you only have 3 factors, both models will fit the same—the structural model is just-identified, and thus the fit of a higher-order factor CANNOT be tested

Structural Model Identification: 2 Factor "Indicators"



Measurement Model for <u>Items</u>:
Possible DF = (12*13) / 2 + 12 = 90
Estimated DF = $8\lambda_i + 12\mu_i + 12\sigma^2_{e_i} = 32$
DF = 90 – 32 = 58 → over-identified

Structural Model for Factors:
Possible DF = (4*5) / 2 + 0 = 10

Estimated DF = $4\lambda_F + 0\sigma^2_F + 1\sigma_{F,F} + 4\sigma^2_{d_F}$
— OR —
Estimated DF = $2\lambda_F + 2\sigma^2_F + 1\sigma_{F,F} + 4\sigma^2_{d_F}$

DF = 10 – 9 = 1 → over-identified

However, this model will not be identified structurally unless there is a non-0 covariance between the higher-order factors

Higher-Order Factor Identification

- Possible structural df depends on # of measurement model **<u>factor variances and covariances</u>** (NOT # items)

  - **2 measurement model factors → Under-identified**
    - They can be correlated, which would be just-identified…
    - Higher-order factor be estimated if both lower-order loadings are equal

  - **3 measurement model factors → Just-identified**
    - They can all be correlated OR a single higher-order factor can be fit
    - Some # variance/disturbances per factor (so, 3 total) in either option
    - Factor variances and covariances will be perfectly reproduced

  - **4 measurement model factors → Can be over-identified**
    - They can all be correlated (6 correlations required; just-identified)
    - They can have a higher-order factor (4 loadings; over-identified)
    - **The fit of the higher-order factor can now be tested**

Examples of Structural Model Hypothesis Testing

- Do I have a higher-order factor of my subscale factors?
  - If 4 or more subscale factors: Compare fit of alternative models
    - Saturated Baseline: All 6 factor covariances estimated freely
      Alternative: 1 higher-order factor instead (so DF=2)—is model fit WORSE?
  - If 3 (or fewer) subscale factors: CANNOT BE DETERMINED
    - Saturated baseline and alternative models will fit equivalently (unless lower-order factor loadings or disturbance variances are constrained to save DF)

- Do I need a residual covariance, but I'm doing IFA in ML?
  - Predict those two items with a factor, fix both loadings=1 if you need a positive covariance or −1/+1 if you need a negative covariance
  - Estimate its factor variance, which then becomes the residual covariance
- Do I have need additional "method factors"?
  - Some examples…

Illustrative Example: "Life Orientation"

Table 2
Means, Standard Deviations, and Correlations for E. C. Chang et al.'s (1994) Life Orientation Test Data

Item	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Item 1	1.00						
Item 2	.51	1.00					
Item 3	.44	.53	1.00				
Item 4	−.16	−.22	−.26	1.00			
Item 5	−.28	−.38	−.33	.50	1.00		
Item 6	−.24	−.29	−.30	.51	.70	1.00	
Item 7	−.22	−.35	−.30	.44	.54	.52	1.00
M	2.24	2.40	2.56	1.85	1.39	1.32	1.40
SD	1.00	0.99	0.99	1.05	1.03	1.00	1.07
Skewness	−0.12	−0.35	−0.57	0.25	0.63	0.68	0.71
Kurtosis	−0.65	−0.36	−0.11	−0.72	−0.14	0.01	−0.23

Note. $N = 389$.

Maydeu-Olivares & Coffman (Psychologicial Methods, 2006) present 4 models by which to measure a latent factor of optimism using the 3 positively and 4 negatively worded items shown below

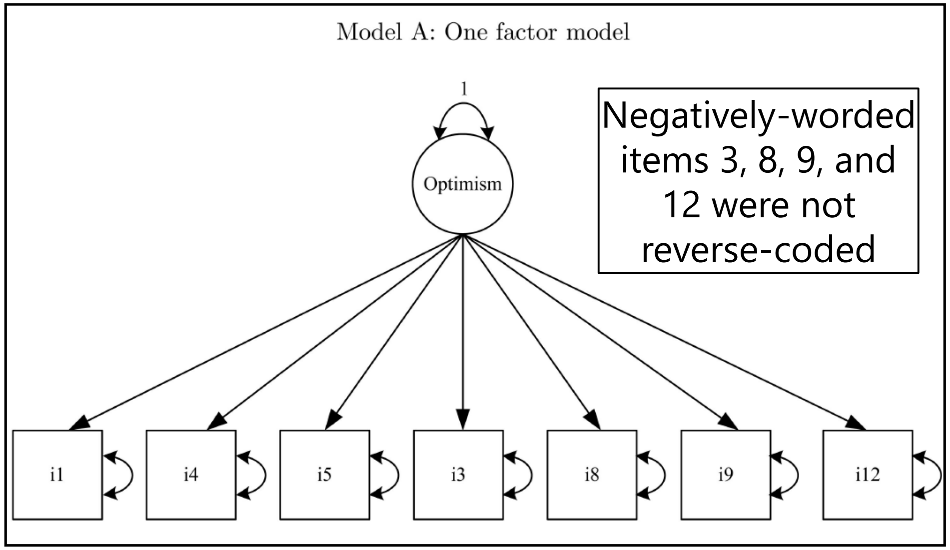
Table 1
Life Orientation Test (LOT) Items (E. C. Chang et al., 1994)

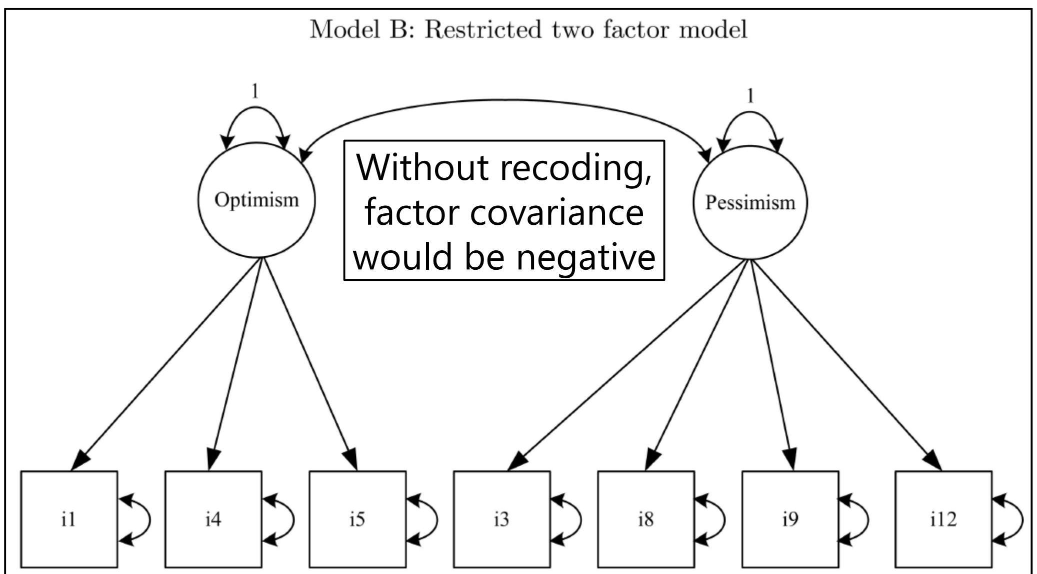
Item	Original item number
1. In uncertain times, I usually expect the best. (positive)	Item 1
2. I always look on the bright side of things. (positive)	Item 4
3. I'm always optimistic about my future. (positive)	Item 5
4. If something can go wrong for me, it will. (negative)	Item 3
5. I hardly ever expect things to go my way. (negative)	Item 8
6. Things never work out the way I want them to. (negative)	Item 9
7. I rarely count on good things happening to me. (negative)	Item 12

Note. The original item number is the order in which the item appears on the actual LOT questionnaire.

- A: Single factor (DF = 14)
- B: Two wording factors (DF = 13)
- C: Three-factor "Bifactor" model (DF = 7)
- D: "Random Intercept" 2-factor model (DF = 13)

What to do with method effects?





Maydeu-Olivares & Coffman (2006) present 4 ways to measure a latent factor of optimism with 3 positively and 4 negatively worded items

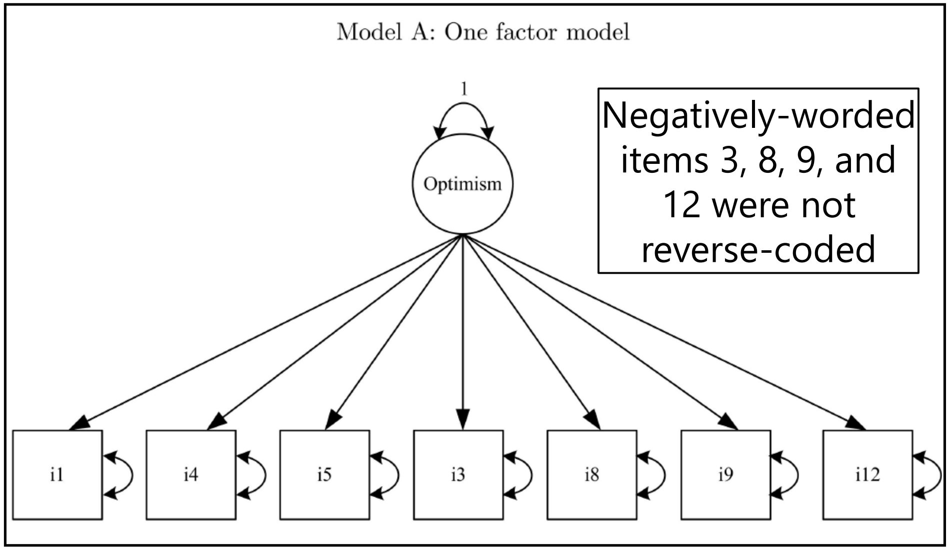
A: Single "optimism" factor (which doesn't fit well)

```
Opt BY i1* i4* i5*
       i3* i8* i9* i12*;
Opt@1; [Opt@0];
```

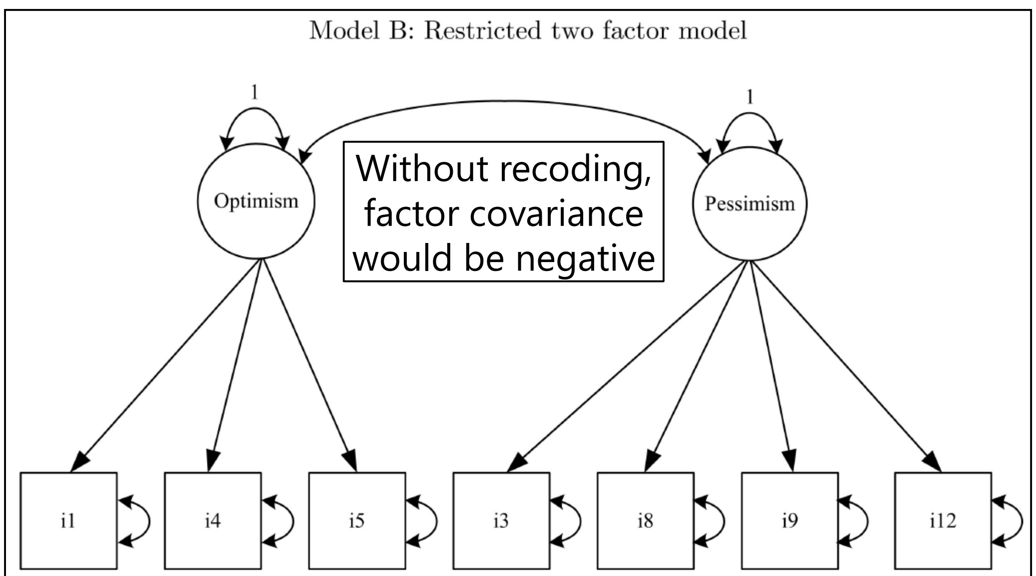
B: "Optimism" and "Pessimism" two-factor model (fits better)

```
Opt BY i1* i4* i5*;
Pes BY i3* i8* i9* i12*;
Opt WITH Pes*;
Opt@1; [Opt@0];
Pes@1; [Pes@0];
```

One- vs. Two-Factor Models

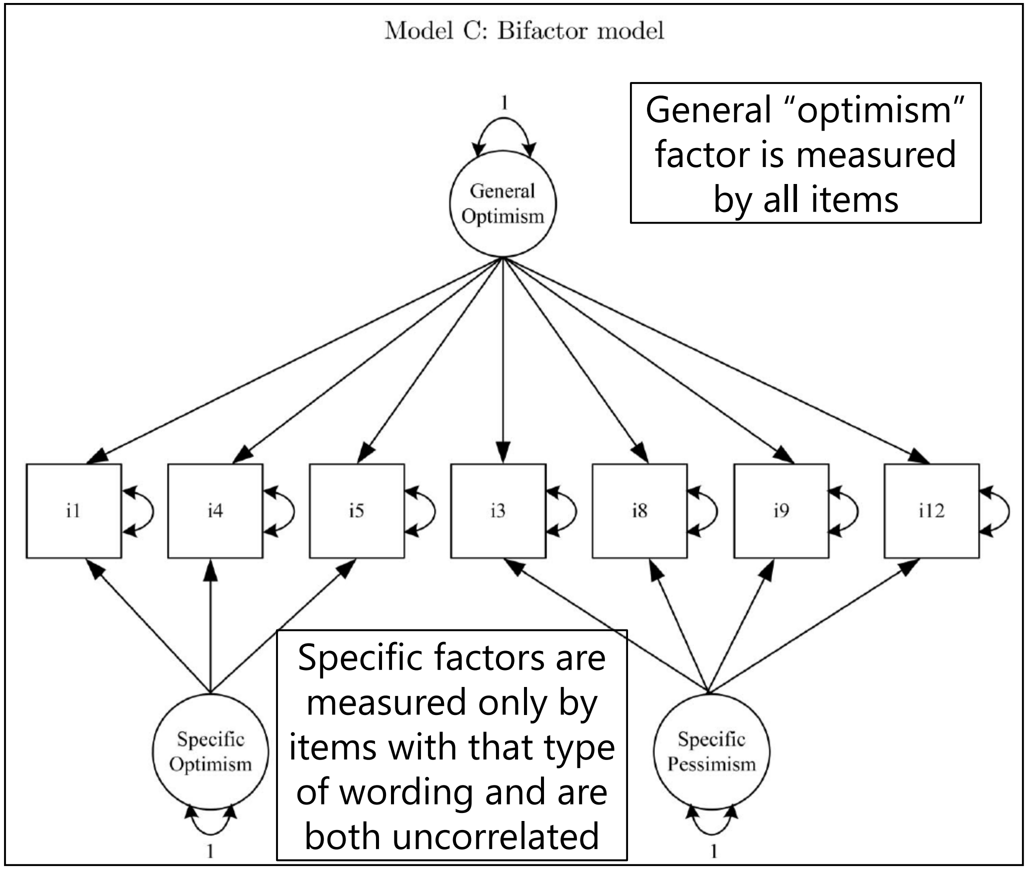


Item	One-factor model: Optimism	Two-factor model Optimism	Two-factor model Pessimism
Item 1	0.38 (0.05)	0.64 (0.05)	0
Item 2	0.48 (0.05)	0.78 (0.05)	0
Item 3	0.46 (0.05)	0.68 (0.05)	0
Item 4	−0.64 (0.05)	0	0.65 (0.05)
Item 5	−0.86 (0.05)	0	0.87 0.05)
Item 6	−0.79 (0.05)	0	0.82 (0.05)
Item 7	−0.70 (0.05)	0	0.70 (0.05)



Note: a higher-order factor could be included if both loadings were fixed to 1, but it would fit the same as just allowing the two wording factors to covary.

Bifactor Model C Fits Well…




```
 Gen BY i1* i4* i5*
        i3* i8* i9* i12*;
 Opt BY i1* i4* i5*;
 Pes BY i3* i8* i9* i12*;
 Gen@1; Opt@1; Pes@1;
[Gen@0  Opt@0  Pes@0];
 Gen WITH Opt@0 Pes@0;
 Opt WITH Pes@0;
```

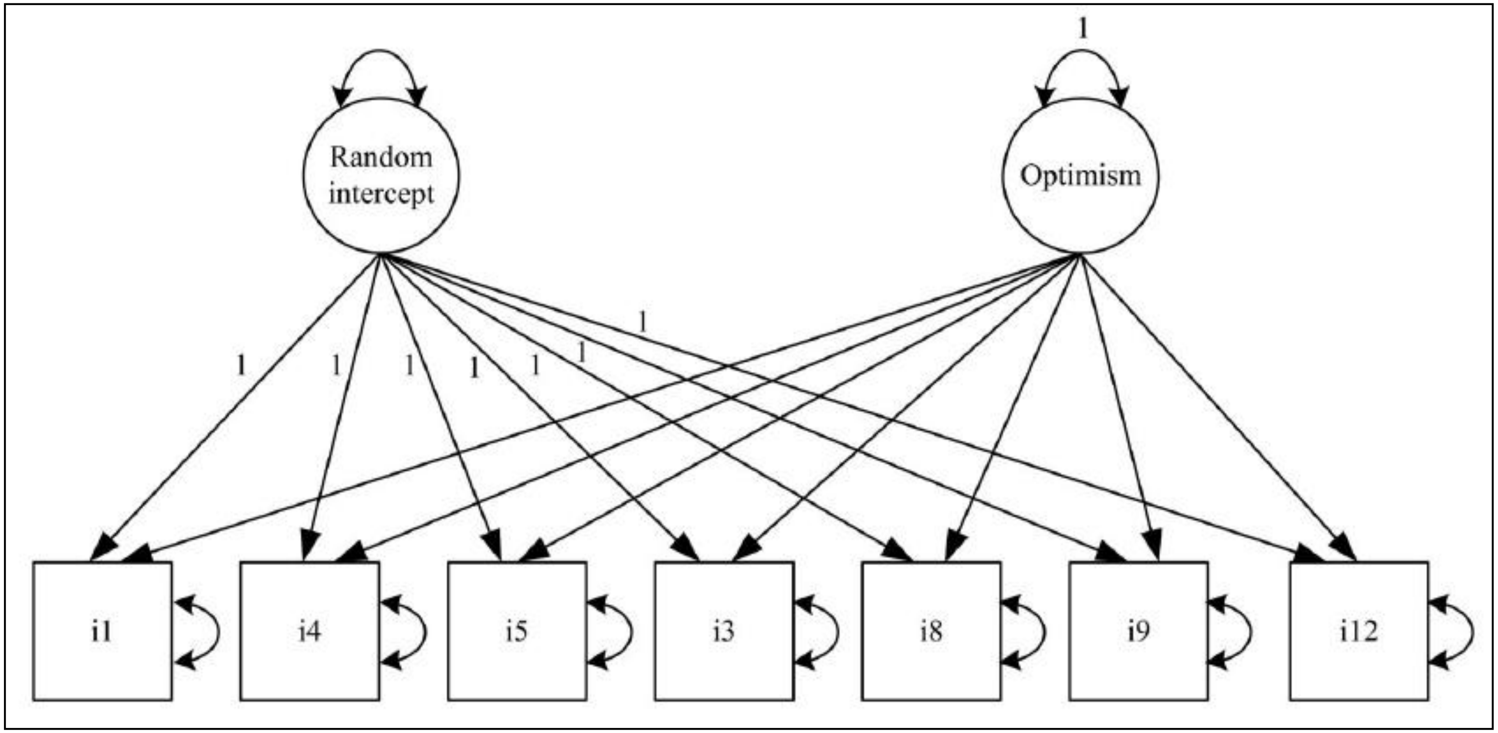
Bifactor model		
Overall optimism	Specific optimism	Specific pessimism
0.35	0.56	0
(0.07)	(0.07)	
0.49	0.61	0
(0.08)	(0.07)	
0.44	0.51	0
(0.07)	(0.07)	
−0.59	0	0.26[a]
(0.09)		(0.18)
−0.76	0	0.38
(0.10)		(0.23)
−0.63	0	0.64[a]
(0.11)		(0.16)
−0.73	0	0.15[a]
(0.08)		(0.18)

2 problems in interpreting these factors as desired:
1) "Specific" positive loadings > "general" loadings
2) Specific negative loadings are weak or non-significant (indicating model is over-parameterized)

Random Intercept Factor Fits Well…

- General "optimism" factor is measured by all items (all loadings estimated)
- New "random intercept" factor allows for constant person shifts across items (e.g., due to different response scale interpretations); Variance = 0.13 here

```
Opt BY i1* i4* i5*
       i3* i8* i9* i12*;
Opt@1; [Opt@0];
Int BY i1@1 i4@1 i5@1
       i3@1 i8@1 i9@1 i12@1;
Int*; [Int@0];
Opt WITH Int@0;
```



One-factor random intercept: Optimism
0.54
(0.05)
0.66
(0.05)
0.61
(0.05)
−0.56
(0.05)
−0.78
(0.05)
−0.71
(0.05)
−0.65
(0.05)

Heartland Forgiveness Scale (HFS)

Yamhure Thompson, L., Snyder, C.R., **Hoffman, L.,** Michael, S.T., Rasmussen, H.N., Billings, L.S., et al. (2005). Dispositional forgiveness of self, others, and situations. *Journal of Personality, 73*(2), 313-360.

Model 4. Six correlated lower-order factors for positive and negative self, other, and situation "forgiveness" and "not unforgiveness" (reverse-coded)

Total possible df for 18 items = 189

$$\frac{v*(v+1)}{2}+v=\frac{18*19}{2}+18=189$$

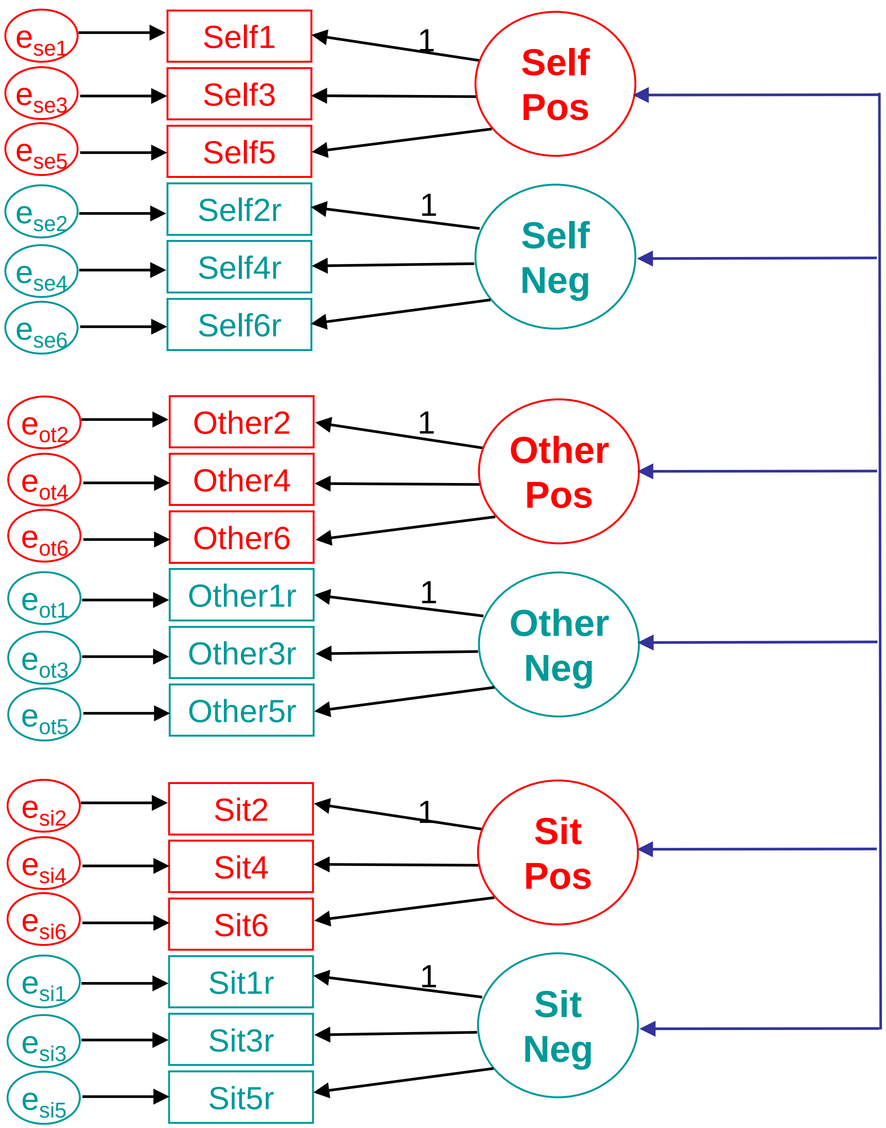
Measurement Model = 48 parameters

$12\lambda_i + 18\mu_i + 18\sigma^2_{e_i}$

Structural Model = 21 parameters

$6\sigma^2_F$, 15 factor covariances (all possible, abbreviated with arrows from line)

Total model DF = 189 – 69 = 120

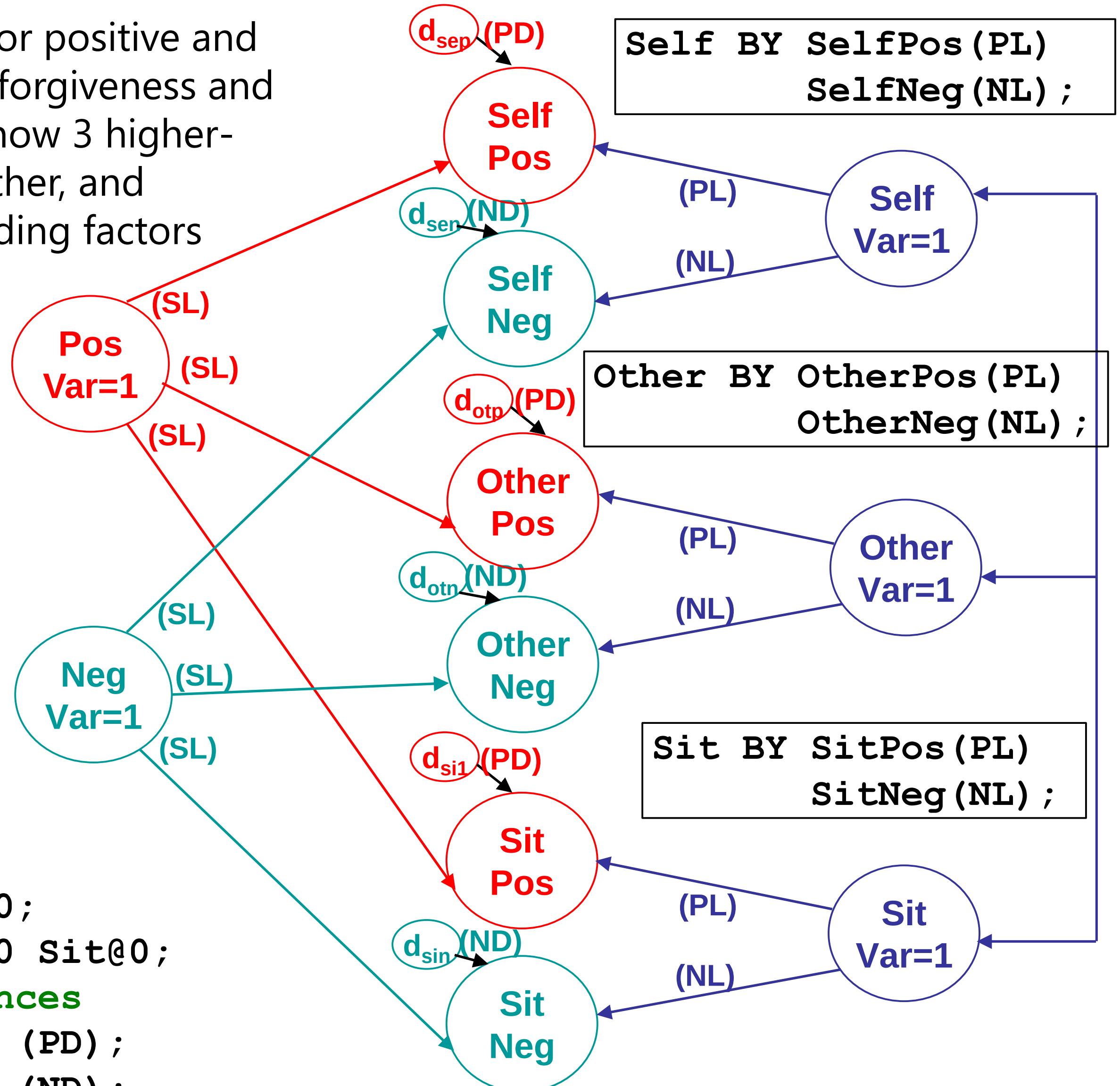


HFS Structural Model

Model 5. Six lower-order factors for positive and negative self, other, and situation forgiveness and not unforgiveness as before, but now 3 higher-order correlated factors of Self, Other, and Situation, and 2 uncorrelated wording factors

Structural Model = 8 parms
$(\mathbf{DF} = \mathbf{21} - \mathbf{8} = \mathbf{13})$

```
! Constant Method Effects
Pos BY SelfPos*   (SL)
       OtherPos*  (SL)
       SitPos*    (SL);
Neg BY SelfNeg*   (SL)
       OtherNeg*  (SL)
       SitNeg*    (SL);
! No method factor covs
Self@1 Other@1 Sit@1;
Self  WITH Other* Sit*;
Other WITH Sit*;
Pos@1 Neg@1; Pos WITH Neg@0;
Pos Neg WITH Self@0 Other@0 Sit@0;
! Constant factor disturbances
SelfPos* OtherPos* SitPos* (PD);
SelfNeg* OtherNeg* SitNeg* (ND);
```

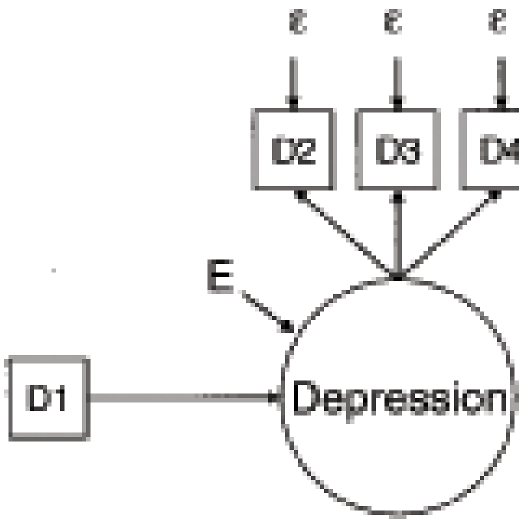


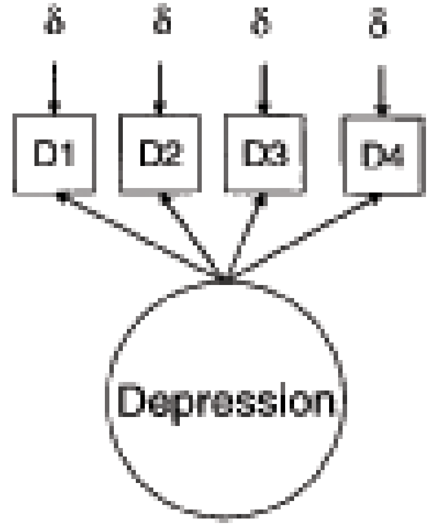
Equivalency across Models

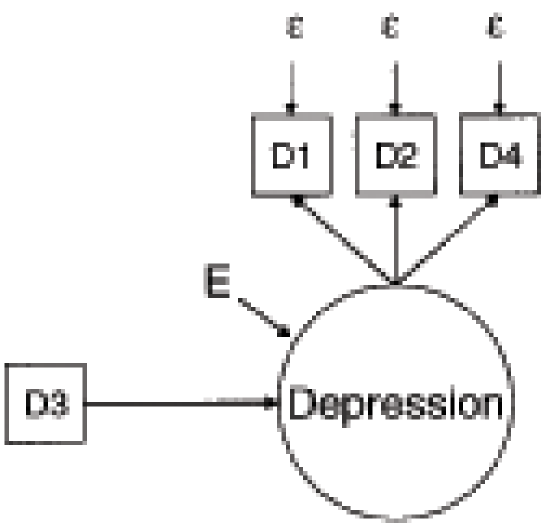
- Remember, the purpose of a measurement model is to reproduce the observed variances, covariances, and means of the items
- This means that models that generate the same predicted variances, covariances, and means of the items are equivalent models
- This will often not be comforting, but it is the truth…
- Here's an example: These models make very different theoretical statements, but they will nevertheless fit equivalently

- Generally speaking, the fewer DF left over (i.e., the more complicated the model), the more equivalent alternative solutions there are

More Equivalent Models…







Top: One can think these 4 items as "effects" (indicators) of depression…

Left: One can think of any one item as "causing" depression and the others as "effects" of depression…

Point of the story: CFA/SEM cannot give you TRUTH. Contrary to what it's often called, SEM is not really "causal" modeling

Wrapping Up…

- Fitting measurement and structural models are two separate issues:
  - **Measurement model**: Do my lower-order factor loadings predict the *observed covariances among my ITEMS*?
  - **Structural model**: Do higher-order factor loadings predict the *estimated covariances among my measurement model FACTORS/THETAS*?
    - A higher-order factor is NOT the same thing as a total score, but it is a way to rescue a multidimensional trait that you want to think of as unidimensional in how it relates to other constructs (i.e., those relations can be specified with just higher-order factor)
- Figure out the measurement models FIRST, then structural models
  - I recommend fitting measurement models separately per factor, then bringing them together once you have the items for each factor/theta fitting well
  - This will help to pinpoint the source of misfit in complex models
- Keep in mind that structural models may not be 'unique'
  - Mathematically equivalent models can make very different theoretical statements, so there's no real way to choose between them if so…